

COFFEE SHOP SALES ANALYSIS

Submitted By

Penmetsa Navya

navyapenmetsa94@gmail.com

Table of Contents

1. Introduction
2. Objective
3. Dataset Overview
4. Data Pre-processing
5. Data Transformation
6. Visualization & Insights
7. Outcomes & Recommendations
8. Conclusion

1. Introduction

In today's competitive coffee shop industry, data-driven decision-making is crucial for optimizing sales, understanding customer behaviour, and enhancing store performance. Analysing sales trends, peak transaction times, best-selling products, and profit margins helps businesses make informed strategic choices.

This project focuses on analysing coffee shop sales, product performance, and revenue trends through data pre-processing, feature engineering, and Power BI reporting. The goal is to extract meaningful insights, detect patterns, and drive business growth through data analytics.

2. Objective

The objective of this project is to analyse sales transactions of a coffee shop, identify trends, and generate insights that can support decision-making. The key goals include:

- a. Sales Performance Analysis: Understanding revenue trends and sales patterns.
- b. Customer Behaviour Analysis: Identifying peak transaction times and popular products.
- c. Store Location Insights: Analysing performance across different store locations.
- d. Product Performance Evaluation: Determining best-selling and underperforming products.
- e. Revenue Optimization: Identifying opportunities to increase revenue through data-driven decisions.

3. Dataset Overview

The dataset contains various columns providing essential sales information, including:

- a. Transaction ID: Unique identifier for each sale.
- b. Transaction Date & Time: Timestamp of the sale.
- c. Store ID & Location: Identifies which store made the sale.
- d. Product Details: Category, type, and product name.
- e. Unit Price & Quantity: Determines revenue per transaction.
- f. Total Revenue: Revenue generated from each transaction.

4. Data Pre-processing

- a. Loading Data: The dataset was initially imported from an SQL database and processed using Python.

- b. Handling Missing Values: No missing values were found in key revenue-related columns.
- c. Data Type Conversion:
 - i. transaction_date converted to Date Time format.
 - ii. transaction_time extracted and converted to Time format.
 - iii. Hour extracted from transaction_time for peak hour analysis.
- d. Duplicate Handling: No duplicate records were found in the dataset.
- e. Outlier Detection: Boxplots were used to detect outliers in Unit Price and Quantity Sold. Outliers were not removed as they were justified by business context.
- f. Feature Engineering: To enhance data insights custom columns like 'Profit Margin %' and 'Is_Peak_Hour' column was created. 'Quantities' column was discretized into Small, Medium, Large for better analysis. 'product type' column was frequency encoded to represent categorical importance.

5. Data Transformation

Power BI Measures Created:

- a. Total Revenue:
SUM(coffeeshop_processed [Total Revenue])
- b. Best-Selling Product:
TOPN(1, SUMMARIZE(coffeeshop_processed, coffeeshop_processed [product_detail], "Revenue", SUM(coffeeshop_processed [Total Revenue])), [Revenue], DESC)
- c. Underperforming Product:
TOPN(1, SUMMARIZE(coffeeshop_processed, coffeeshop_processed product_detail], "Revenue", SUM(coffeeshop_processed [Total Revenue])), [Revenue], ASC)
- d. Avg Products Sold Per Day:
*VAR TotalProducts = DISTINCTCOUNT(coffeeshop_processed[product_id])
VAR TotalDays = DISTINCTCOUNT(coffeeshop_processed[transaction_date])
RETURN DIVIDE(TotalProducts, TotalDays)*
- e. High Revenue Transactions:
COUNTROWS(FILTER(coffeeshop_processed, coffeeshop_processed[total_revenue]>AVERAGE(coffeeshop_processed[total_revenue])))
- f. Transactions Per Hour:
COUNT(coffeeshop_processed[transaction_id])

6. Visualizations

- a. Revenue Trend Analysis (Line Chart) - Displays revenue trends over time with drill-down options.

- b. Top 5 Product Category (Bar Chart) – Displays top 5 product category by revenue.
- c. Product Performance (Tree map) - Highlights category wise product performance with total revenue.
- d. Peak Hour Sales (Clustered Column Chart) - Identifies high-traffic transaction times and their revenue impact.
- e. Store Performance (Funnel Chart) - Compares profit margin % across different store locations.
- f. Sale by Store (Pie Chart) – Compares sales and transactions across different store locations.
- g. Quantity by Product (Matrix) – Shows number of Quantity Size sold per product with background conditional formatting.
- h. Best Selling & Underperforming Product (Card) – Displays Best & Worst Selling product by revenue.
- i. KPI Cards: Display High Revenue Transactions, Product Performance, and Average Products Sold per Day.
- j. Slicers: Filter by Store Location, and Quantity Size.
- k. Buttons: Page number, bookmarks and blank buttons are used to make the dashboard interactive.

Each visualization provides valuable insights into sales patterns, operational efficiency, and revenue opportunities.

7. Outcomes & Recommendations

- a. Key Insights:
 - i. Peak Sales Hours: Transactions peak during morning (8 AM - 10 AM) indicating key revenue period.
 - ii. Best-Selling Products: Coffee Beans and Specialty Drinks generate the most revenue across multiple stores.
 - iii. Underperforming Products: Branded Merchandise and Accessories contribute the least to sales.
 - iv. Quarter & Month Peak Sales: Quarter 2 and months May & June show high sales.
 - v. Profit Margins: Certain products show lower profit margins affecting overall profitability.
- b. Recommendations:
 - i. Increase Staffing During Peak Hours and Months to reduce wait times and improve customer experience.
 - ii. Prioritize Promotions on Best-Selling Products to maximize revenue growth.
 - iii. Reduce Discounts on High Demand Products to protect profit margins.
 - iv. Improve Placement of Underperforming Products to increase visibility and sales.

8. Conclusion

This coffee shop sales analysis successfully pre-processed and transformed sales data to extract actionable insights. Through Power BI dashboards, peak sales hours, best-selling products, store performance variations, and revenue optimization opportunities were identified. These insights can be leveraged to enhance sales strategies, improve store efficiency, and maximize profitability.

Future work could involve predictive modelling to forecast seasonal demand and customer segmentation for personalized marketing campaigns.